



Fickle product mix: exporters adapting their product vectors across markets

Lionel Fontagné, Angelo Secchi, Chiara Tomasi

► To cite this version:

Lionel Fontagné, Angelo Secchi, Chiara Tomasi. Fickle product mix: exporters adapting their product vectors across markets. 2014. hal-01299822

HAL Id: hal-01299822

<https://hal.science/hal-01299822>

Preprint submitted on 8 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUTE
OF ECONOMICS



Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

WORKING PAPER SERIES

Fickle product mix: exporters adapting their product vectors across markets

Lionel Fontagné §
Angelo Secchi *
Chiara Tomasi °

§ University of Paris 1 and CEPII

* University of Paris 1 and LEM-Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy

° University of Trento, Italy, and LEM-Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy

2014/26

December 2014

ISSN (ONLINE) 2284-0400

Fickle product mix: exporters adapting their product vectors across markets *

Lionel Fontagné[†], Angelo Secchi[°], and Chiara Tomasi[‡]

[†]PSE - University of Paris 1 and CEPPI

[°]PSE - University of Paris 1 and LEM-Scuola Sant'Anna Pisa

[‡]University of Trento, LEM-Scuola Sant'Anna Pisa and Banque de France

First Draft
December 2014

Abstract

This paper analyzes how multi-product firms adjust their exported product-mix across destinations. Using cross sections of Italian and French data, we show that firms do not follow a rigid ordering in their product mix exported in different markets but rather they adapt their choices to better match with country characteristics. By using metrics based on export shares and on sequences of product names we provide new insights on the extent a firm's products portfolio changes across destinations that go beyond simple rank correlations. Demand asymmetries, market structure heterogeneity and differential abilities to match unit values of products supplied by competitors emerge as three significant factors in explaining the variety-country variability observed in firms' export patterns. Our results resist when we control for a firm's choice of not exporting an available product to a given destination, an explicit choice likely to contain relevant information.

JEL codes: F14, L11, L22

Keywords: multi-product, multi-country firms, product vectors, demand and concentration

*The present work has been possible thanks to a research agreement between the Italian Statistical Office (ISTAT) and the Scuola Superiore Sant'Anna. Lionel Fontagné acknowledges support by the Bank of France (Competitiveness and International Trade Analysis Division of the Microeconomic and Structural Analysis Directorate). Angelo Secchi gratefully acknowledges the Paris School of Economics for granting him a 'Residence de Recherche' for the period 2012-2013. We also acknowledge financial support from the Institute for New Economic Thinking, INET inaugural grant #220.

1 Introduction

Recent evidence based on transaction level data shows that an overwhelming share of international trade is conducted by large firms that export a broad variety of products to different destinations.¹ In the US firms shipping more than five (HS10) products² represent 30% of total exporters but account for 97% of all exports (Bernard et al., 2009), in Brasil 25% of exporters ship ten or more (HS6) products and account for 75% of total exports (Arkolakis and Muendler, 2010). Similar figures are observed for Italy and France where 42% and 40% of exporters ship more than five (HS6) products and account for 96% and 95% of total export flows, respectively.

While these multi-product firms typically concentrate their sales in a few products (or “core products”³) that generate the majority of firms’ exports, we tend to observe a great deal of variation in the set of products exported across destination markets. Indeed, a firm’s product mix, i.e. the set of products exported by a firm to a country, is rather fickle across countries. Two interesting patterns emerge with respect to a firm’s product mix variability. Take as an illustrative example a firm producing electrical motors and shipping 16 different products to 20 different destinations.⁴ First, in 4 out of the top 10 destinations⁵ (DEU,FRA,GRC,NLD) this firm exports the same product mix composed by “AC Motors” of two different powers combined with an additional generic good labeled “Parts of these motors”. In other 3 of the top 10 destinations (AUS, USA, CZE) sales are concentrated on a different product mix based on “Universal AC/DC motors. Second, the relative importance of the 3 products exported in the EU top destinations, while stable in a first approximation, change across markets. Indeed, this firm is exporting to Germany mainly high power AC motors (those with power ranging between 750W and 75KW) while in Netherlands the most important products in term of export share are the low-power AC motors (those with power below 750W).⁶

To what extent are these patterns common among exporters and what are the main factors driving the variation of a firm’s product mix choice across destination markets? Using a cross section of firm-product-destination level Italian and French data, the paper provides an exhaustive investigation of export patterns of multi-product firms across destination markets. The present work contributes to the existing literature by providing robust evidence that firms adjust their exported product-mix across destinations and shedding light on the main drivers behind this adaptation mechanism. Indeed, a rigorous analysis of the export patterns of multi-product firms requires both

¹See Bernard et al. (2011); Mayer et al. (2014); Eckel et al. (2011).

²The Harmonized Commodity Description and Coding System, also known as the Harmonized System (HS) nomenclature, is an internationally standardized system of names and numbers to classify traded products. See Section 2 for more details.

³The micro-trade literature on multi-products exporters usually defined as core products as the top-selling varieties for which a firm has some systematic sales advantage. More generally, the business literature defines as “core” or “flagship” products those goods which are most directly related to a company’s core competencies.

⁴This is a real example. We indeed do not unveil which is the firm and omit, on purpose, some information for confidentiality reason. A detailed description of the data-set used for the empirical investigation is provided in Section 2.

⁵The top 10 destinations cover 90% of the total exports of this firm.

⁶Figure A1 in Appendix A shows the export sales distribution across products and destinations for this firm.

the identification of salient empirical facts on the firms’ product sales distribution across countries, and the establishment of the determinants of such patterns. Our paper takes on both tasks.

First, the paper provides a robust set of empirical regularities concerning firms’ product mix choices across markets. In studying product hierarchies across countries we move beyond simple rank correlations. We use a dissimilarity index based on product market shares which allows us to better capture the economic importance of each product in terms of export shares. Further, we add an original investigation of the extent a firm’s products portfolio (i.e. the set of products exported) changes across destinations by using sequences of product names. All together the results suggest that there is a strong variation observed in a firm-product exports across countries. Firms do not follow a rigid ordering in their product mix across markets but rather they adjust their choices in each destination market. The results resist when we controls for a firm’s choice of not exporting an available product to a given destination, an explicit likely to contain relevant information.

Second, the paper offers possible explanations for the observed variability in the export patterns of multi-product firms. While firm-product heterogeneity in efficiency is an important determinant of firms’ entry into export markets, a regularity largely confirmed in empirical studies, it can not explain the variability in a firm’s product-mix across destination. Demand asymmetries, aiming to capture tastes for different varieties, market structure heterogeneity, which proxies for the level of competition, and market positioning of a firm with respect to its competitors emerge as three significant factors in explaining such variability.

The remainder of the paper is organized as follows. The next section provides an overview of the data used in the empirical analysis. Section 3 motivates our paper by presenting some key regularities on multi-product and multi-country firms which have been useful in disciplines models of international trade and by adding new stylized facts that are difficult to reconcile with the existing theoretical frameworks. Section 4 presents the empirical results on firms’ product shares across destinations while Section 5 focuses on firms’ product vectors. The last section concludes.

2 Data

Our analysis on multi-product and multi-country firms exploits trade data on the universe of Italian and French exporters. The foreign trade statistics data-set are collected by the national statistical office and consist of all cross-border transactions performed by Italian and French firms, respectively.⁷ For all export flows defined at the firm-product-destination level the data include both annual values and quantities expressed respectively in euros and in kilograms. Product categories are classified according to the Harmonized System and they are reported at the 6-digit level (HS6).⁸ We define a variety as a firm-product combination. Since we are interested in the cross-section of firm-product exports across destinations, we restrict our attention on a single year, for 2006.⁹

⁷The datasets were accessed at the ISTAT and Banque de France facilities and they have been made available for work after careful screening to avoid disclosure of individual information.

⁸Although French custom data are available at a finer level of aggregation (CN8), we consider the classification at 6-digit to make the analysis comparable across the two countries.

⁹This is the last year available for the Italian dataset.

Table 1: DESCRIPTIVE STATISTICS, 2006

	Italy		France	
	Whole sample (1)	Restricted sample (2)	Whole sample (3)	Restricted sample (4)
Total exports (billion euros)	271.1	266.2	268.7	265.1
# Firms	74,365	45,530	32,432	19,689
Avg. # Products	8.4	12.6	8.5	12.9
Avg. # Countries	9.5	14.7	8.8	13.5

Notes: Table reports descriptive statistics on the Italian and French dataset for 2006. In the restricted sample we keep those firms exporting more than one product and serving more than one destination.

Table 2 reports the total value of exports, the number of firms, the average number of products and destinations for 2006 for the two countries. Columns 1 and 3 show some descriptive statistics for the whole sample of Italian and French manufacturing firms, respectively.¹⁰ Since our study focuses on the variability of a firm’s exports across products and countries we exclude from the analysis those firms exporting only one product and serving only one destination. As reported in columns 2 and 4, while the reduction in the number of firms is sizable the amount of exports covered remains substantially unchanged.¹¹

3 Motivating the paper

Recent advances in research on multi-product and multi-country exporters provide a set of new empirical regularities on both the intensive and extensive margins of trade at the firm-product level within and across destinations. Arkolakis and Muendler (2013, 2010) report a set of three robust empirical stylized facts for Brasil, Denmark, Chile and Norway characterizing an exporter’s number of products shipped within each destination (*export scope*) and the corresponding average sales per product within each destination (*export scale*).¹² First, few large wide-scope exporters and many narrow-scope firms coexist within each destination. Second, the sales of wide-scope exporters are concentrated in few products and the same firms are able to tolerate lower sales for their lowest-selling products. Third, within each market the mean (across firms) exporter scope and the mean (across firms) exporter scale are positively associated.

¹⁰Table B1 in the Appendix B reports the total value of exports and the number of exporters distinguishing between three broad categories of firms: manufacturers, wholesalers and retailers, others. As clear in the Table, in Italy there is a larger fraction of manufacturing exporters than in France (55% vs 32%). These firms account for a large share of a country’s total exports (85% in Italy, 71% in France). On the contrary, France is characterized by a larger number of wholesaler and retailer exporters than Italy (43% in France against 36% in Italy.). Wholesalers and retailers account for 21% of France’s total exports but only 12% in Italy.

¹¹Because of some missing value in the employment variable we are working with a slightly smaller number of firms in the case of France (19,043 rather than 19,689).

¹²Exporters’ scope and scale are defined here following Arkolakis and Muendler (2010).

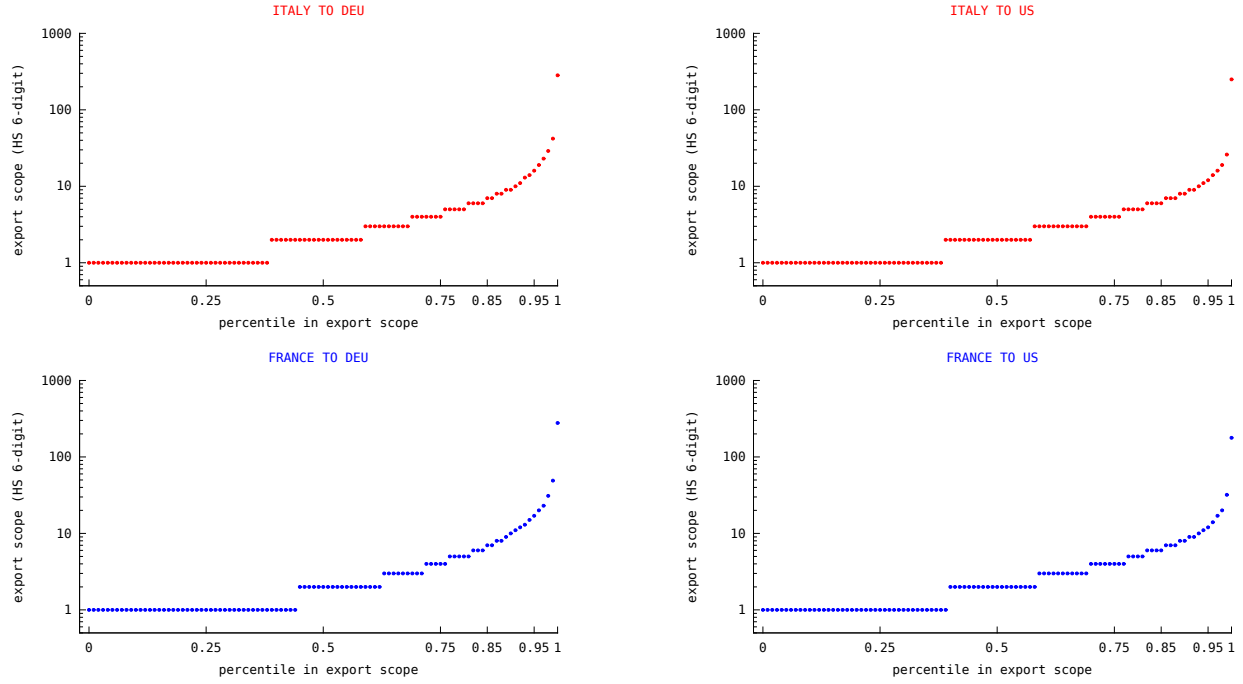


Figure 1: Italian (**Top panels**) and French (**Bottom panels**) exporter scope distribution in Germany (left panel) and in the USA (right panel). Custom data for 2006. Products at H6 6-digit level. Logarithm scale on the vertical axis.

Adding upon these evidences, the same regularities appear to characterize both Italian and French exporters. Figure 1 reports the relationship, for Italian firms (top panel) and French firms (bottom panel), between exporter scope and the corresponding percentile in the exporter scope distribution for firms shipping to Germany and to the USA.¹³ The distribution appears very skewed: more than 70% of exporters export only 3 or less products while only the top 10% ship more than 10 products.

In Figure 2 we plot, for firms exporting 4,8,16, and 32 products, the average (across firms) product export value for products sharing the same rank against their rank within firm. These figures confirm firstly that also in Italy and France wide-scope exporters are indeed much larger, in terms of sales, than narrow-scope ones. But more importantly, wide-scope exporters are to tolerate lower sales of their lowest selling products. With this respect, regressing the exporter's lowest ranked product (log) sales against its (log) exporter scope in a market conditioning on fixed effects for firm and destination returns an elasticity of -1.45 (0.007) and -1.56 (0.011) for Italy and France, respectively.¹⁴

Finally, Figure 3 confirms also for Italian and French firms that within destination mean exporter scope and the corresponding mean exporter scale are positively associated.

This set of stylized facts concerning export patterns' of multi-product and multi-country firms

¹³Similar patterns are found for other destinations.

¹⁴Arkolakakis and Muendler (2010) estimate an elasticity for Brazilian firms of 2.1.

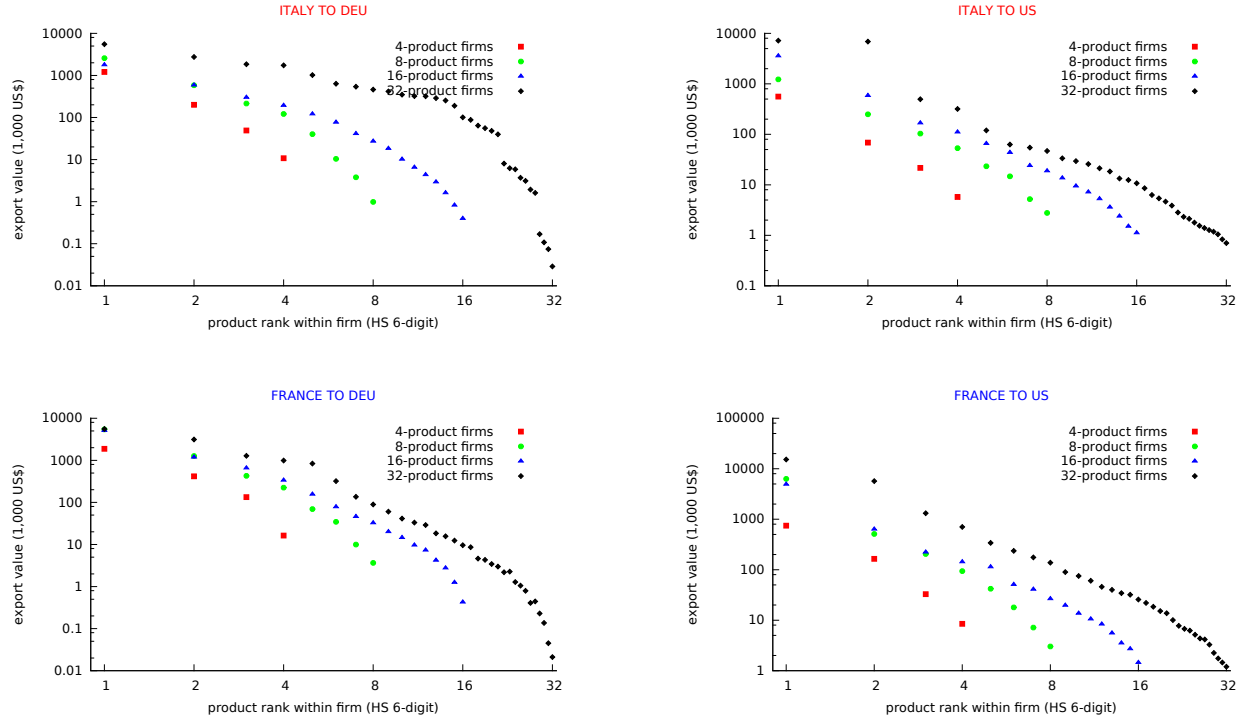


Figure 2: Italian (**Top panels**) and French (**Bottom panels**) firms' product sales distributions in Germany (left panel) and in the USA (right panel). The figure reports the distribution for firms exporting exactly 4,8,16 or 32 products. Custom data for 2006. Products at H6 6-digit level. Logarithm scale on both the vertical and the horizontal axis.

has opened new avenues for theory. Two additional ingredients are, indeed, required in a standard Melitz type framework to match this set of stylized facts: a product-destination specific incremental local entry costs and a firm-product efficiency both declining with a firm's scope.¹⁵ Arkolakis and Muendler (2010) show that if the firms' combined costs, the incremental local entry cost component and the marginal cost schedule, are strictly increasing in a firm's scope, then sales within a destination are concentrated in a few core products, wide-scope exporters sell more of their top selling products than narrow-scope exporters and sales of the lowest selling products decline with scope. At the same time, the model can also generate the observed positive relationship between mean exporter scope and scale and explain the high frequency of narrow-scope exporters.

The second of these two assumptions, that is firm-product efficiency declining with scope, implies that firms enter export markets with the most efficient product first and then expand their scope moving-up the marginal-cost ladder. Then, successful products in one destination should also be leading products in all the other markets and the same hierarchy of products should be observed across different destinations. The available evidence on this point suggests, however, that, even if the correlation among firm-product sales rankings is positive, we are far from observing a perfectly

¹⁵Other models considering heterogeneity in firm-product efficiency are in Eckel and Neary (2010); Eckel et al. (2011); Bernard et al. (2011); Mayer et al. (2014).

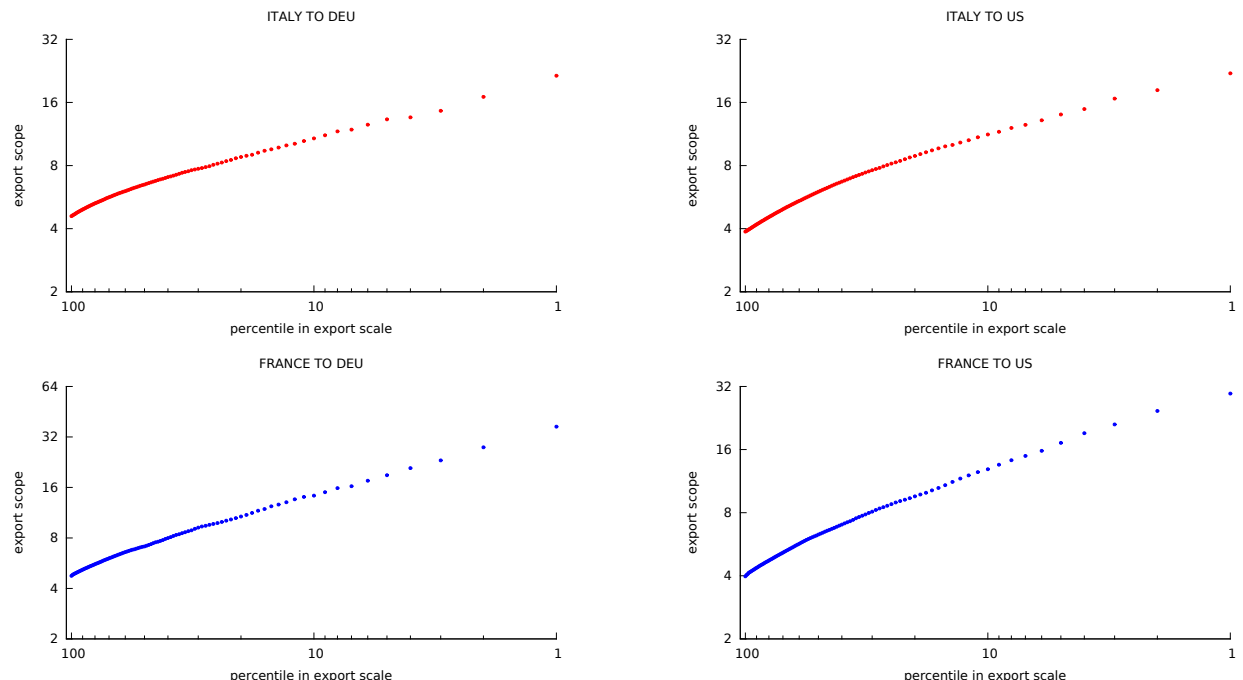


Figure 3: Relationship between mean Italian (**Top panels**) and French (**Bottom panels**) exporter scale and mean exporter scope in Germany (left panel) and in the USA (right panel). Custom data for 2006. Products at H6 6-digit level. Logarithm scale on both the vertical and the horizontal axis.

stable hierarchy across destinations.

Using data on Brazilian firms, Arkolakis and Muendler (2010) compute the correlation based on firm-product sales between the ranking in a destination, using either US or Argentina as reference country, and the ranking elsewhere, obtaining on average coefficient of 0.85. Based on French data, Mayer et al. (2014) define the global and local rankings using the value of a firm-product exports to the world and to a single destination, respectively, and find a correlation value of 0.68. For a set of countries Di Comite et al. (2014) consider the pairwise bilateral correlation for each country-pair of firm-product Belgian export quantities, resulting in a measure of 0.50.¹⁶

Existing models, predicting either perfectly stable (Arkolakis and Muendler, 2010; Mayer et al., 2014) or totally uncorrelated (Bernard et al., 2011) orderings of products across destinations, are not consistent with the observed evidence.¹⁷ The reason of this inconsistency is likely to reside in the absence, in these theoretical frameworks, of a source of heterogeneity that is variety-country specific. Di Comite et al. (2014) is the only attempt, we are aware of, to overcome this limitation. In fact, in their model consumer preferences are asymmetric across varieties and heterogeneous across markets. However, they do not consider multi-product firms, but rather single variety companies.

¹⁶Similar conclusions have been reached by empirical analyses that look at the firm sales variation in different countries (Eaton et al., 2011; Kee and Krishna, 2008).

¹⁷Mayer et al. (2014) emphasizes that the rigid ordering in a firm's product ranking across destinations predicted by their model does not perfectly hold empirically.

Table 2: R^2 FOR REGRESSIONS ON VARIETY-COUNTRY (LOG) EXPORT SALES.

	HS 6-digit			CN 8-digit		
Firm-Product FE	Yes		Yes	Yes		Yes
Country-Product FE		Yes	Yes		Yes	Yes
ITALY	48.9%	22.9%	57.1%			
FRANCE	53.0%	22.8%	61.2%	53.5%	23.0%	61.7%

Notes: Results are based on restricted sample imposing a minimum of two destinations for each firm-product and of two firms for each country-product. Italian and French custom data for 2006.

Before proceeding we conduct, in the next section, a preliminary exercise to assess the existence and, eventually, to quantify this variety-country specific heterogeneity component.

3.1 The variety-country heterogeneity component

The variety-country heterogeneity component can be captured empirically by regressing the export value of each transaction (that is a firm-product-country observation) on a set of fixed effects. Specifically we estimate the following regression model

$$y_{fpd} = \alpha_{fp} + \alpha_{pd} + \epsilon_{fpd}, \quad (1)$$

where y_{fpd} are firm-product-destination export sales, α_{fp} is a firm-product fixed effect that captures all the variety-specific factors, such as cost or quality, that explain the variation in the data, α_{pd} is a product-country fixed effect aiming to capture all the country characteristics, which are product specific, and ϵ_{fpd} that is the variety-country heterogeneity we aim to quantify. To get some variability in the data we need to restrict our analysis to those products exported by each firm to more than one country and those country-product combinations served by more than one firm. This allows us to identify a variety and country-product specific effect in the regressions.

Results are reported in Table 3.1. They confirm the relative importance of firm-product heterogeneity in explaining the variation in the data. Indeed, firm-product FE explain about 50% of the exports variability. This justifies theoretical approaches that allow efficiency varying by firm-product rather than by firm only.¹⁸ At the same time regressions with only country-product fixed-effects suggests that part of the variability is driven by products' determinants that varies across countries. However, what is most important here is that all the models leave a substantial amount of variability unexplained. When we combine firm-product and country-product FE the R^2 is between 57% and 61% for Italy and France respectively, leaving more than one third of the variability unexplained. Thus, the data suggest that a non-negligible part of the sales variation

¹⁸Both Di Comite et al. (2014) and Munch and X. (2008) reach similar conclusions, pointing out that the firm-product dimension is relevant in explaining the data variation.

is due to variety-specific heterogeneity. The existence of this heterogeneity implies that the rank correlation of a variety sales across destinations needs not to be equal to one.¹⁹

In what follows we provide an exhaustive investigation of multi-country and multi-product firms export with the twofold aim of clarifying the sources the variety-country variability we observe in the data and of identifying other factors that, in addition to production and entry costs, can help explaining it.

4 Firms' product shares across destinations

This section documents stylized facts about the variation in firm-product exports sales across trade-partner countries. We first examine the correlation of firm-product export sales ranking among countries, by overcoming two limitations faced by previous empirical analyses. First, we control for a firm's choice of not exporting an available product to a given destination, an explicit likely to contain relevant information. Second, we implement a more rigorous dissimilarity measure. Finally, we empirically identify some of the factors that help to explain the variety-country variability in firms' exports across destinations.

4.1 A firm's products market shares across destinations

As detailed in Section 3, theoretical models with multi-product and multi-country firms generate a rigid ordering within a firm regarding the products exported across markets. In these models firms face incrementally higher marginal costs of production for those varieties that are far away from the core competency. As market share is increasing in productivity, which is firm-product specific, a firm that sells to different countries should have product market shares that are strongly positively correlated across its destinations. However, the previous section shows that empirically we observe a variation in the export sales of identical varieties sold in different countries. This implies that the theoretical prediction for stable variety market shares rankings across destinations does not necessarily hold empirically. In this section we investigate the pairwise correlations of exporters' product market shares across destinations.

We identify, for each firm, the local product shares vector which is the vector containing the export share of each product (HS6) exported by firm f to destination c . We call this *Local Product Shares Vector* (LSPV $_{fd}$). While the dimension of the local product shares vector of a given firm may differ across destinations, the sum of the components of each LSPV $_{fd}$ is always equals to one. We then compute the pairwise rank correlations between each LSPV $_{fd}$ and all the other destinations served and we take their median value.

Results, reported in Figure 4, confirm the absence of a stable hierarchy of products across destinations. On average the median Spearman pairwise rank correlation is 0.511 and 0.480 for

¹⁹In order to check that our results are not merely driven by a classification effect, we replicated the analysis for French firms by narrowing the definition of a variety and using the firm-CN8 combination. Independent of the product-market definition, we see that a large amount of the variability is left unexplained when we include both firm-product and country-product fixed effects.

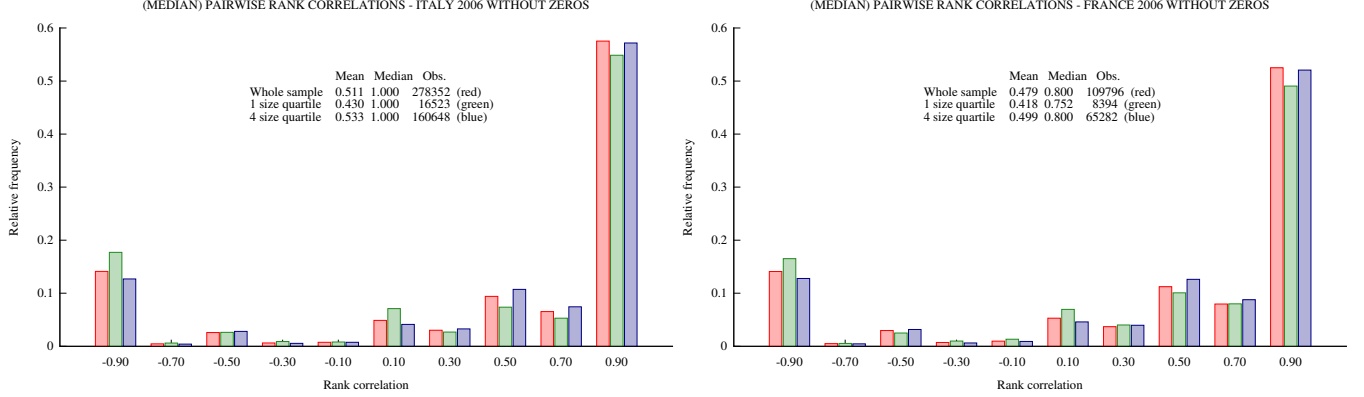


Figure 4: Median pairwise rank correlation for the whole sample (Red), for small firms (Green) and big firms (Blue). Italian firms (**Left panel**) and French firms (**Right panel**). Custom data for 2006.

Italy and France, respectively. This value is very close to that reported for other European countries (Mayer et al., 2014; Di Comite et al., 2014). In Figure 4 we also check whether this correlation is driven by the existence of a large number of small firms that export few products to few destinations, for which local vectors might be very similar among them. Thus, we replicate the analysis for small and big companies defined as those firms belonging to the first and fourth quartile of the employment distribution of our sample. The lack of a perfect hierarchy is confirmed across different size classes.²⁰

There are two potential pitfalls in using rank correlations to assess the degree of heterogeneity of a firm's local product share vector in different destinations: first, the absence of any direct control for a firm's decision of not exporting a product to a given destination, that is product selection and, second, the lack of a proper quantification of the compositional dissimilarity between two different vectors. In the following we consider deeply both issues.

Concerning the product selection, the fact that a firm may decide not to export one of its products to a given destination affects the observed rank correlation. Indeed, any products that is not exported to a certain country is dropped from the local product share vector of that destination. Consider, for instance, a firm exporting three products to three destinations as reported in the following example

$$LPSV_{f1} = \begin{pmatrix} 0.97 \\ 0.02 \\ 0.01 \end{pmatrix} \quad LPSV_{f2} = \begin{pmatrix} 0.99 \\ 0 \\ 0.01 \end{pmatrix} \quad LPSV_{f3} = \begin{pmatrix} 0.51 \\ 0.49 \\ 0 \end{pmatrix} .$$

Without considering the zeros the median pairwise rank correlation is one, while with zeros

²⁰In unreported figures, available upon request, we replicate the analysis by gradually restricting the sample to those products and countries that represent more than 1% or 5% of a firm's total exports. These cut hardly changes the pairwise rank correlations.

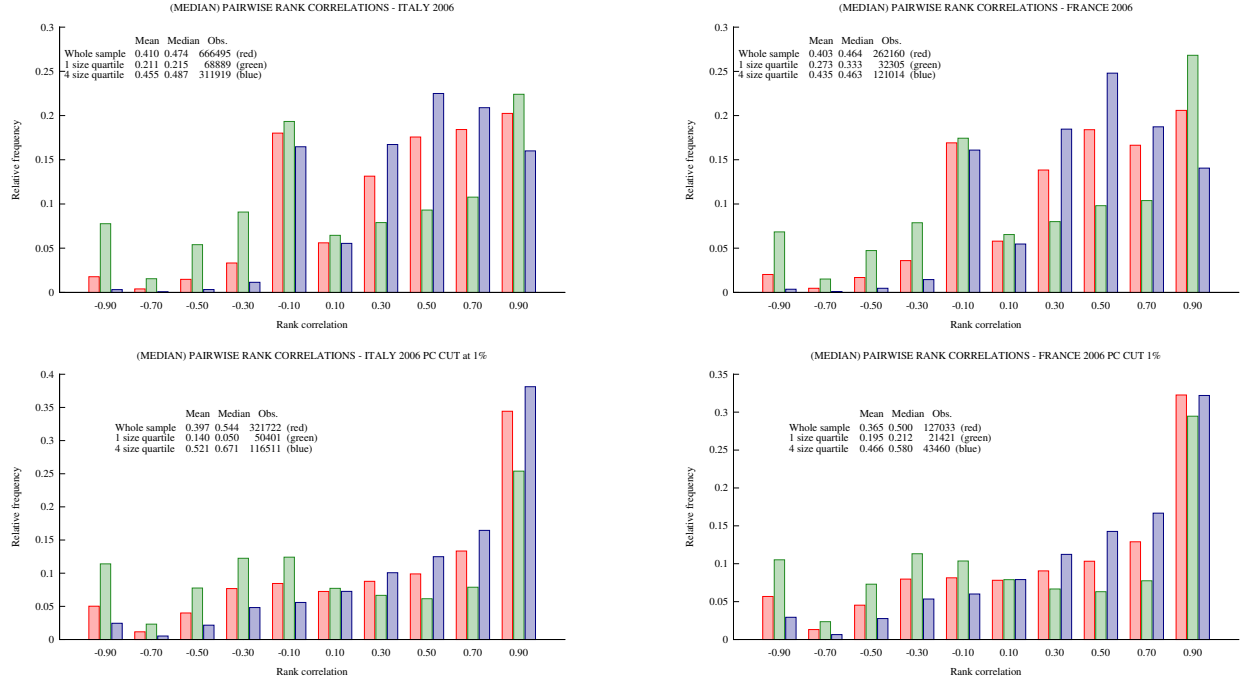


Figure 5: Median pairwise rank correlation for the whole sample (Red), for small firms (Green) and big firms (Blue), controlling for product selection. *Top panel:* all firm's products and countries. *Bottom panel:* products and countries accounting for more than 1% of a firm's total exports. Italian firms (**Left panel**) and French firms (**Right panel**). Custom data for 2006.

the correlation reduces to 0.50. To overcome this limitation we control for product selection by filling with a zero share any product that is not exported to a destination. As a consequence all the $LSPV_{fd}$ turn out to be of the same dimension.

In the top panel of Figure 5 we replicate the analysis of the pairwise Spearman rank correlation, directly controlling for a firm's decision of not exporting a product to a given destination. The changes in the shape of the distribution is quite evident: on average the median pairwise rank correlation decreases to 0.410 for Italy and to 0.403 for France. Similar reduction is reported for the sub-samples of small and big firms. We conclude that, when controlling for the extensive margin response, there is a substantial departures from a steady order of export products for exporters across destinations. This result suggests that the country-specific demand for each variety might offer an explanation not only for the variation in a firm-product intensive margin, but also for a firm's product extensive margin across destinations.

The strong reduction in the rank correlation might be partly driven by the over-inflation of zeros associated to "irrelevant" products and destinations. Therefore, as a consistently check, we investigate whether the inclusions of those products not exported to a destination may have affected our results. We do so by looking only at a set most relevant products and markets, that are those representing more than 1% of a firm's total exports. The bottom panel of Figure 5 reveals that the

correlation is quite stable: on average the value is indeed 0.397 and 0.365 for Italy and France.²¹

The second limitation in the use of the Spearman correlation concerns the fact that the coefficient is based on ranked variables and it does not take into account the effective “dissimilarity” between a firm’s product market shares in different countries. Indeed, rank correlation cannot capture any heterogeneity in a situation like the following

$$LPSV_{f1} = \begin{pmatrix} 0.97 \\ 0.02 \\ 0.01 \end{pmatrix} \quad LPSV_{f3} = \begin{pmatrix} 0.51 \\ 0.49 \\ 0 \end{pmatrix}$$

where the pairwise rank correlations is equal to one, although notable differences are detected by $LPSV_{f1}$ and $LPSV_{f3}$. To overcome this limitation we use the Bray-Curtis dissimilarity measure, a statistic that allows us to quantify the compositional dissimilarity between two different vectors. This dissimilarity index is defined as

$$BC_{fdj} = \frac{\sum_p |s_{fdp} - s_{fjp}|}{\sum_p |s_{fdp} + s_{fjp}|} \quad , \quad (2)$$

where s_{fdp} and s_{fjp} represent export shares of product p in destinations d and j . $|s_{fdp} + s_{fjp}|$ is a normalizing factor (equal to 2 in this case). The Bray-Curtis dissimilarity is bound between 0 and 1, where 0 means the two vectors have the same composition, i.e. high correlation, ($s_{fdj} = s_{fjp}$ for all p) and 1 means the two vectors are completely disjoint, i.e. no correlation. In the example above the BC dissimilarity computed between $LPSV_{f1}$ and $LPSV_{f13}$ equals 0.47.

In Figure 6 we therefore report the pairwise Bray-Curtis dissimilarity computed between the local vectors considering all the products and countries (top panel) and after excluding those products and destinations accounting for less than 1% of a firm’s total exports (bottom panel). The analysis confirms that there is a substantial deal of dissimilarity: between 30% and 40% of firms have median pairwise Bray-Curtis distance close to 1. This result provides further support to the existence of a variety-country specific form of heterogeneity that explain part of the variability observed in the firm-product composition and exports across destinations. Again, this result is not totally driven by the over-inflation of zeros associated to “irrelevant” products or countries, as documented in the bottom panel of Figure 6 where we exclude marginals goods and destinations. Results are consistent for both countries of origin.

4.2 The determinants of dissimilarity between local vectors

The analysis so far reveals a substantial departure from a perfect ordering of the products exported by a firm in different destinations. At the same time, the pattern observed is at odds with a scenario where products are randomly assigned across countries. What are the main factors explaining firms’ products choices across countries? What are the determinants of the dissimilarity between a firm’s

²¹A similar value is observed when excluding products and markets that account for less than 5% of a firm’s total exports.

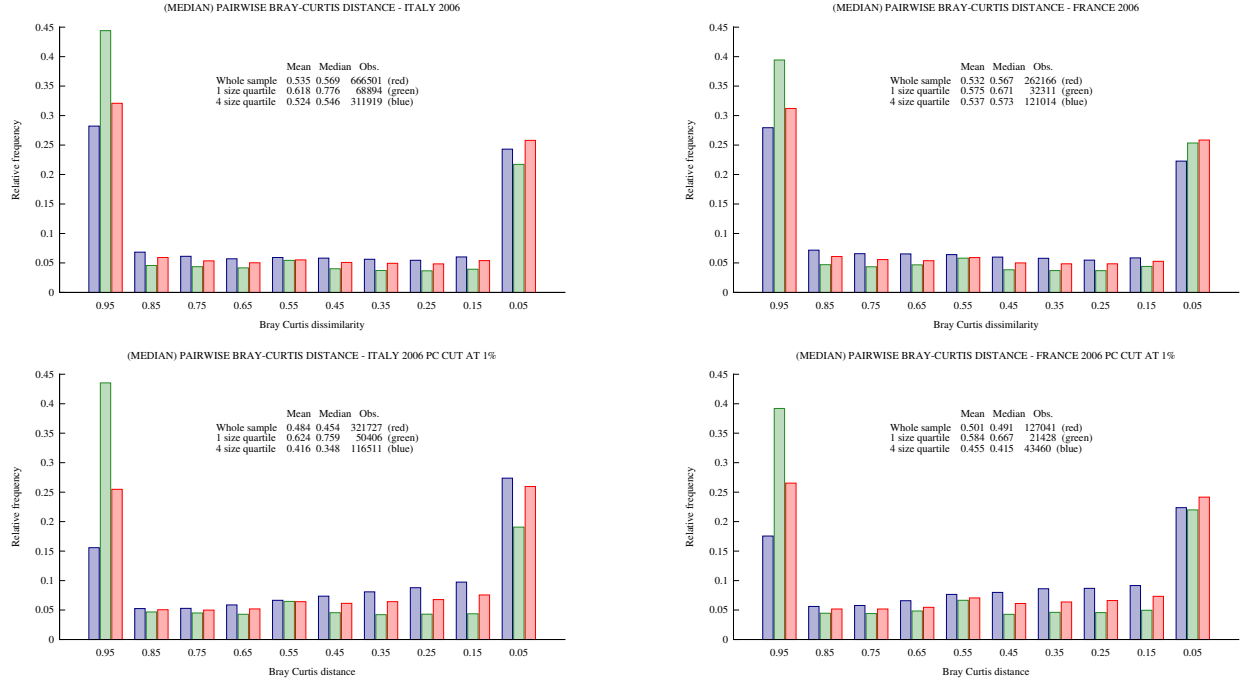


Figure 6: Median pairwise Bray-Curtis dissimilarity index for the whole sample (Red), for small firms (Green) and big firms (Blue), controlling for product selection. *Top panel:* all firm's products and countries. *Bottom panel:* products and countries accounting for more than 1% of a firm's total exports. Italian firms (**Left panel**) and French firms (**Right panel**). Custom data for 2006.

local product share vectors?

Theoretical models of multi-product and multi-country firms based on cost efficiency as the only determinant of firm behavior in export markets cannot account for all the variation detected in the data. Indeed, although firm-product heterogeneity in costs has empirically been confirmed to be very important in determining firms' behavior into export markets, it seems to be not a sufficient condition to explain the firm-product level sales variation in different countries. In the following we propose three factors likely to help in explaining, together with firm heterogeneity, the observed firm-destination variation.

First, competition effects in a destination market are likely to influence the choice of the product mix exported to that market. In Mayer et al. (2014), for example, a firm responds to tougher competition by dropping its worst performing products and by concentrating its export sales towards the best performing ones. The variation in competition, however, does not alter the rigid ordering of a firm's product ranking but only the product mix skewness across destinations.

Second, demand heterogeneity, driven by asymmetric preferences between varieties across countries, is likely to explain why a firm's product market share and rankings vary among destinations. In this respect, the model proposed by Di Comite et al. (2014) in the context of single-product firms by allowing each destination country to have different preferences for the same variety offered, shows that a firm's sales vary across markets.

A third explanation for the variation in a firm-product export sales, not explicitly considered in any theoretical model, is provided by the difference across countries in the market positioning of a firm with respect to its competitors. Indeed, a firm's choice regarding its exports across countries can be related to how different is the price of the product exported to a specific country with respect to the average price of its competitors in that product-destination.²²

In investigating which factors might help to explain the variety-country variability in a firm's exports, we focus on three firm-country specific determinants: *demand*, *competition* and *market positioning*. We start by calculating a measure of demand computed at firm-destination level. As a proxy for the demand faced by an Italian or French firm f in destination country d we compute the (log) total imports for product p in d , excluding imports from Italy or France, by using product-level (HS6 digit) trade data from BACI (Gaulier and Zignago, 2010).²³ We then weight the total imports for each product-destination by the relative importance of product p in a firm's total export sales so to obtain a firm-destination level demand measure. To mitigate endogeneity problem we use all the set of products exported by the firm rather than the subset of products exported by the firm to a specific destination.²⁴ However, since firms' global export decisions are made jointly with their market specific strategies, we will interpret the results from our specifications in terms of correlations and not causality. The firm-destination level demand measure is given by

$$\text{Demand level}_{fd} = \sum_{p \in \Omega_f} \left(\frac{\text{Exp}_{fp}}{\text{Exp}_f} \right) (\log \text{Imp}_{pd}^*)$$

where Ω_f is the set of products firm f is exporting (to any country), Exp_{fp} and Exp_f are a firm's exports in product p and a firm's total exports, respectively, Imp_{pd}^* is the sum of imports from all countries of origin (except Italy or France) ($\text{Imp}_{pd}^* = \sum_{o \in O_{pd}} \text{Imp}_{pod}^*$). The higher the level of the demand in destination d for a firm f , the stronger the match between the products offered by this firm and the goods demanded in that country.

We then proceed by defining a firm-destination level measure of competition. To measure the level of competition faced by each firm in a given market we again use the BACI dataset. For each of the products belonging to the set of products exported by firm f (in any country) we compute the Herfindahl Index (HHI_{pd}^*) in the destination country d . The product-destination level of competition is given by the negative of the log HHI weighted, as before, by the relative importance of product p in a firm's total exports

$$\text{Competition}_{fd} = \sum_{p \in \Omega_f} \left(\frac{\text{Exp}_{fp}}{\text{Exp}_f} \right) (-\log \text{HHI}_{pd}^*) .$$

²²Price variation across firms and destinations can be driven by several factors such as quality differences, markups heterogeneity, market competition, firm composition, supply factors such as shipping costs, and further destination country characteristics. The purpose here is not to evaluate the sources of within-firm price adjustments across destinations but rather to measure the market positioning, based on price differences, of a firm with respect to its competitors.

²³The variable obtained using the BACI dataset are signed with *.

²⁴A similar approach has been adopted in Bernard et al. (2014).

where the HHI_{pd}^* is computed using as shares the imports of destination d in product p from origin o over the total imports of country d in product p from all the origins (excluding Italy or France), that is $(Imp_{pdo}^* / \sum_{o \in O_{pd}} Imp_{pdo}^*)$. The higher the value of the variable, the stronger is the level of competition faced by a firm in a foreign market.

Finally, we compute a measure of market positioning for each firm f in destination d as the negative of the logarithm absolute value difference between the price of a firm's product p to destination d and the average price charged by the other countries of origin for the same product p in country d

$$\text{Market Positioning}_{fd} = \sum_{p \in \Omega_f} \left(\frac{Exp_{fp}}{Exp_f} \right) (-\log | UV_{fpd} - UV_{pd}^* |) .$$

where UV_{fpd} is a firm's product-country unit value obtained as the ratio of export values to export quantities and UV_{pd}^* is the weighted geometric average of the unit values (from BACI) in the product-destination from different origins but not Italy or France. Precisely, it is computed as $UV_{pd}^* = \exp \left(\sum_{o \in O_{pd}} \left(\frac{Imp_{pod}^*}{Imp_{pd}^*} \right) \log UV_{pod} \right)$, where the unit value of product p in destination d from country o (UV_{pod}) is weighted by the relative importance of that transaction on the total imports in product-destination $(\frac{Imp_{pod}^*}{Imp_{pd}^*})$. The higher the value of this index, the closer is the market positioning of a firm with respect to its competitors.

Using these measures, we estimate the following regression model

$$BC_{fd} = \beta X_{fd} + \alpha_f + \alpha_d + \epsilon_{fd} \quad (3)$$

where we regress the firm-country Bray-Curtis dissimilarity index (BC_{fd}) on a vector, X_{fd} , with the three firm-country specific variables: a proxy for a firm's demand in destination d , a proxy for the competition level a firm's face in d and the market positioning of the firm in the same market. Since we seek to uncover the variation in the dissimilarity index for a given firm, we include in the model firm fixed effects throughout (α_f). These fixed effects account for systematic differences across exporters in ability that might affect trade outcome across destinations. We also include destination fixed effects, α_d , which implicitly account for cross-country differences in total income, market toughness, trade costs and other institutional frictions that might affect the variation in a firm's local product vector shares measured by the BC index. We are thus interested in β which reflects the sign of the conditional correlation between the dissimilarity index and our firm-destination level variables. Note that since on the right-hand side of our regression we have the logarithm of a weighted geometric mean we can interpret the estimated coefficient as semi-elasticities. The error term, ϵ_{fd} , includes possible other firm-destination idiosyncratic factors explaining the dissimilarity index. Indeed, while the variety-specific country differences proposed here offer a plausible explanations for the dissimilarity between a firm's local product share vectors across destinations, alternative explanations may also be at work. Standard errors are clustered at firm and country level.

Table 3: REGRESSIONS ON THE BRAY-CURTIS DISSIMILARITY MEASURE

Dependent variable: BC_{fd}					
	Benchmark	1% cut	5% cut	Fraction model with CRE	
	(1a)	(1b) ^a	(2) ^b	(3) ^b	(4) ^c
<i>ITALY</i>					
Demand level _{fd}	-0.010*** (0.000)	-0.052*** (0.005)	-0.012*** (0.002)	-0.008*** (0.002)	-0.010*** (0.001)
Competition _{fd}	-0.018*** (0.002)	-0.020*** (0.003)	-0.018*** (0.003)	-0.014** (0.003)	-0.003** (0.001)
Market Positioning _{fd}	0.050*** (0.003)	0.205*** (0.009)	0.039*** (0.003)	0.028*** (0.002)	0.054*** (0.001)
N.Obs	661,501	661,501	321,727	133,036	661,501
Adj. R^2	0.68	0.68	0.70	0.72	-
<i>FRANCE</i>					
Demand level _{fd}	-0.008*** (0.001)	-0.052*** (0.007)	-0.008*** (0.002)	-0.008*** (0.002)	-0.003* (0.001)
Competition _{fd}	-0.020*** (0.003)	-0.024*** (0.003)	-0.023*** (0.003)	-0.018** (0.004)	-0.012*** (0.002)
Market Positioning _{fd}	0.047*** (0.003)	0.191*** (0.014)	0.031*** (0.003)	0.024*** (0.003)	0.048*** (0.001)
N.Obs	264,613	264,613	128,390	56,696	264,613
Adj. R^2	0.67	0.67	0.68	0.70	-
Firm FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes

Notes: Table reports regression on the BC dissimilarity measure. Data are for 2006. ^a Column (1b) reports standardized coefficients. ^b In column (2) and (3) marginal products and destinations, those counting less than 1% and 5% of total export respectively, are removed from the sample. ^c In column 4 we estimate a Fractional Probit model with Correlated Random Effects CRE discussed in Papke and Wooldridge (1996). Standard errors clustered at firm and country level.

***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table 3 contains the results. The coefficients reported in Column 1a, which presents the benchmark specification, show a very significant impact of the three firm-country level variables on a firm's dissimilarity index. Once again, we interpret the sign of β 's as the sign of a conditional correlation that does not reflect causality. We observe that the higher the demand level for a firm in a given destination, i.e. the higher the match between products demanded in a destination and a firm's available products, the lower the median dissimilarity of the LPSV in that destination with respect to those exported to all other countries served by the firm. The result indeed suggests that it is more likely for a firm to differentiate less its product portfolio in terms of export shares in those countries where the demand is high for its most important products, i.e. when there is a high "match" between the products supplied by the firm and those demanded in the country. The coefficient is a semi-elasticity and it can be interpreted as the change in the dissimilarity index from a percent increase in the demand level, holding all other variables in the equation constant.

Therefore, the point estimates in column 1 indicates that doubling the demand faced by a firm in destination is associated with a 0.01 decrease in the (median) dissimilarity index for both Italy and France.

Concerning the market-structure measure, we observe that the higher the competition faced by a firm selling its products in a destination market d , the lower the dissimilarity between the LPSV in that country and the other countries j . Thus, we observe lower variability in a firm's export shares with respect to other destinations when exports is directed to markets with tougher competition. Note that this result is at least in part in line with the finding of Mayer et al. (2014). In fact, they observe that a firm responds to tougher competition by dropping its worst performing products and by concentrating its sales towards its better performing goods. We claim here that the level of competition affects the variability of firm-product export sales across destinations by inducing a firm's to differentiate less its product share vector when facing markets characterized by stronger competition. The partial semi-elasticity of the market-structure measure for Italy and France implies that an increase of 100% would induce a drop in the (median) dissimilarity index of 0.018 and 0.020, holding the other variables constant.

The third variables, market positioning, is positive and statistically significant both for Italy and France. This suggests that the dissimilarity index increases in countries where a firm's position in terms of its export prices is in line with that of its competitors. Put differently, the result indicates that a firm is more likely to reallocate its resources on relatively marginal products and diversify more in terms of export share in markets where its prices are similar to that of its competitors. The effect of a firm's price positioning seems to be particularly strong. Indeed, the estimated coefficient indicates that a rise of 100% in the market positioning of a firm increases the dissimilarity by about 0.050 for both Italy and France.

To see which, among the three factors, is more relevant in explaining the firm-product sales variation across countries we computed the standardized regression coefficients, reported in column 1b. The results reveal that all the three regressors have a statistically significant impact on the BC measure. However, the examination between the three variables tells us that market positioning has a stronger effect, while a comparable effect is detected for demand heterogeneity and competition level. If we increase the market positioning by one standard deviation, the (median) BC measure for Italian firms will increase by 0.200 standard deviations while increasing the demand or the concentration by one standard deviation will result in only a 0.052 and 0.020 standard deviations decrease, respectively. It appears that, in terms of standard units, increasing market positioning is more than three times as effective as changing the demand level and about 10 times as effective as increasing the competition in the market. Similar results are observed for France.

In columns 2-4 we conduct a sensitivity analyses to ensure that our findings are not driven by the over-inflation of zeros associated to "irrelevant" products or destinations (columns 2-3) and, second, that our results are not influenced by the fractional nature of the dependent variable (column 4). In column 2-3 we progressively restrict our sample to a firm's products and countries that represent at least 1% and 5% of its total exports. The two specifications in Table 3 confirm the robustness of

our baseline results regarding the relevance of both firm-country specific variables in explaining the variety-country variability of a firm’s exports. In column 4 we keep the full sample of countries and products but re-estimate the model using the Fractional Probit model with Correlated Random Effects CRE discussed in Papke and Wooldridge (1996).

5 Firms’ product mix across destinations

So far we have studied the variation in firm-product exports across destinations by using a firm’s product export shares and by measuring the dissimilarity between its local vectors. However, this analysis provides only an incomplete description of the phenomenon since it forces to identify the “core” products of a firm only on the base of sale values. Because of product complementarity or technological relatedness, we might observe that the production and thus the export of one product by a firm results in the production and the export of components or related goods that are not necessarily characterized by high sales.²⁵ The latter products, although irrelevant from a “quantitative” point of view, are crucial for the definition of a firm’s core competences.

The analysis that follows overcomes this limitation by discarding product export shares and by looking at the names of products exported by each firm across destinations. First, by looking at the frequency at which products are exported across markets we define a firm’s core product vector. Then, we investigate the determinants of a firm’s choice to export its core product vector to a given destination. Finally, we measure how much different is a firm’s local product vector from its core vector and investigate the determinants of such difference.

5.1 Firms’ core product vector

The first issue to be addressed is how to empirically identify a firm’s core product vector. The approach we follow here departs from that of previous section where we consider a firm’s product export share. In the following we define a firm’s product vector simply as a list of product names. We then single out a firm’s core product vector by observing the frequency at which products vectors are observed across markets.

First, we define a firm’s *Local product vector* (LPV_{fd}) which is a binary vector reporting 1 when a product is actually exported by firm f to destination d and 0 otherwise. LPV_{fd} represents one specific arrangement of the $2NP_f$ possible, where NP_f represents the total number of products exported by f .²⁶

Table 4 reports descriptive statistics on firms’ export diversification for Italian (top panel) and French (bottom panel) firms. The first two rows of each panel show the statistics for the number of products exported and the number of destinations served. We observe that the average Italian exporter exports 12 products and it reaches more than 14 countries. Similar findings are observed

²⁵Figure A2 in Appendix A shows the example of a firm exporting to many different countries three high sales products (different types of motors) and one low sales product (parts of motor).

²⁶Figure A3 in Appendix A reports the example of LPV_{fd} for the electric motors producer.

for France, where the average number of products exported and destinations served is 12 and 13, respectively.²⁷

By looking at the LPV_{fd} (third row) we observe that on average firms have around 8 different local product vectors, both in Italy and France. Therefore, firms exhibit a substantial departure from an equable product mix composition across markets. Indeed, the local product vector is neither equals to 1, which would be the case where the same product vector is exported by a firm across destinations, nor equals to the average number of reached destinations, 14 or 13, which would imply a complete differentiation across countries.²⁸

The definition of a firm LPV_{fd} allows us to single out two different types of product vectors: the *Core product vector* (CPV_f) and the *Weak Core product vector* ($WCPV_f$). The former is the most frequent arrangement observed for firm f across markets. Precisely, the CPV_f represents the typical (the mode) product vector exported by firm f .²⁹ The latter is any arrangements which leaves unchanged all the cells with 1 in the CPV_f . That is, the $WCPV_f$ is a product vector containing the CPV_f .³⁰

Last row of each panel of Table 4 reports the statistics for the percentage of destinations to which a firm is exporting the $WCPV_f$. The average fraction, computed across firms, of destinations where the $WCPV_f$ is exported is 0.6 in both countries of origin. Again, this result confirms that although a firm is exporting to more than half of its destinations the core product vector, there is a non negligible percentage of countries where it is selling a different combination of products.

5.2 The choice of exporting the core product vector

Why does a firm decide to export its core product vector in a destination and a different product mix in other countries? What does explain the choice of exporting the core product vector? As shown in the previous sections, a firm's decision about which product to export to a destination depends not only on firm and country characteristics but also on variety-specific idiosyncratic factors. In line with the analysis of Section 4.2 we focus on demand, competition and market positioning differences. Specifically, we propose the following regression model

$$Dwpcv_{fd} = \beta X_{fd} + \alpha_f + \alpha_d + \epsilon_{fd} \quad (4)$$

where $Dwcpv_{fd}$, is a dummy variable equal to 1 if firm f is exporting a $WCPV_f$ in country d , X_{fd} is the vector containing the three firm-country specific variables, a proxy for a firm's demand in a destination, α_f and α_d are firm and country fixed effects, respectively.

²⁷The numbers here refer to the restricted sample where we keep those firms exporting more than one product and serving more than one destination. The same numbers are lower if we consider the whole population of exporters. As shown in Table 2, on average the number of products exported by Italian and French firms is 8, while the number of destinations reached is 9.5 for Italy and 8.8 for France.

²⁸Note that a firm's maximum number of product vectors is the minimum between 2^{NP_f-1} and ND_f , where NP_f is the number of products exported and ND_f the number of destinations served by firm f .

²⁹Regarding the identification of the CPV_f , in case of ties LPV_{fd} are ranked by their value. In other words they are ranked according to the total value exported to destination d .

³⁰Figures A4 and A5 show the CPV_f and the $WCPV_f$ for the electric motors producer.

Table 4: EXPORT DIVERSIFICATION: PRODUCTS, DESTINATIONS and LPV

	ITALY							
	Mean	Std. Dev.	Min.	1Q	2Q	3Q	Max.	Obs.
# Products	12.55	19.18	2.00	3.00	7.00	14.00	555.00	45,455
# Destinations	14.66	15.53	2.00	4.00	9.00	21.00	129.00	45,455
# LPV	8.46	9.80	1.00	3.00	5.00	10.00	117.00	45,455
Share of destination with WCPV	0.62	0.26	0.01	0.44	0.60	0.85	1.00	45,455
Avg. Levenshtein distance	0.44	0.28	0.00	0.21	0.39	0.61	1.00	43,299
	FRANCE							
	Mean	Std. Dev.	Min.	1Q	2Q	3Q	Max.	Obs.
# Products	12.87	21.39	2.00	3.00	6.00	14.00	649.00	19,364
# Destinations	13.54	15.50	2.00	3.00	8.00	18.00	165.00	19,364
# LPV	7.99	9.87	1.00	2.00	4.00	9.00	134.00	19,364
Share of destination with WCPV	0.63	0.26	0.02	0.47	0.62	0.85	1.00	19,364
Avg. Levenshtein distance	0.44	0.27	0.00	0.22	0.39	0.60	1.00	19,364

Notes: Table reports the descriptive statistics for firms' export diversification for 2006. The statistics are computed on the restricted sample, i.e. we remove from the dataset exporters with either a single product or a single destination.

The results, reported in Table 5, show a very significant impact of demand, competition and market positioning on the probability of exporting the weak core product vector. In the initial specification, column 1, we estimate a linear probability model. The coefficient on the demand variable is positive and statistically significant suggesting that the higher the level of demand in a destination, the higher the likelihood that a firm will be able to export its core products in that destination. Precisely, the point estimates indicates that, on average, a 100% percent increase in firm-destination demand is associated with a 1.5 for Italy and 1.9 percentage points for France increase in the probability of exporting the WCPV. Given that the average fraction of destination with $WCPV_f$ is 60%, this corresponds to an increase of around 2.5 percent. This result confirms what we found studying product vectors in terms of export shares and suggests that whenever a firm meets higher demand for its available products, the probability of exporting a product vector that contains its core products increases.

Concerning the competition variable, the results again largely confirm previous conclusions. Tougher competition across export market destinations does indeed have an impact on a firm's exported product mix. The higher the competition level faced by a firm in a destination, the lower the probability a firm will move away from its core productive capabilities and "test" marginal products in that market. As postulated by Mayer et al. (2014) this result may be driven by the fact that firms respond to increases in competition in their market by exporting those products which are most directly related to their core competencies. A 100% rise in market competition

Table 5: REGRESSIONS ON THE PROBABILITY OF EXPORTING THE WEAK CORE PRODUCT VECTOR

Dependent variable: Dwcpv _{fd}				
	Benchmark (1)	1% cut (2) ^a	5% cut (3) ^a	Probit with CRE (4) ^b
<i>ITALY</i>				
Demand level _{fd}	0.015*** (0.001)	0.015*** (0.002)	0.013*** (0.003)	0.019*** (0.001)
Competition _{fd}	0.022*** (0.003)	0.028*** (0.004)	0.026*** (0.005)	0.016*** (0.002)
Market Positioning _{fd}	-0.093*** (0.004)	-0.081*** (0.004)	-0.087*** (0.003)	-0.102*** (0.001)
N.Obs	666,501	321,727	133,036	666,501
Adj. R ²	0.37	0.33	0.12	-
<i>FRANCE</i>				
Demand level _{fd}	0.019*** (0.001)	0.016*** (0.002)	0.017*** (0.004)	0.018*** (0.001)
Competition _{fd}	0.031*** (0.003)	0.038*** (0.006)	0.036*** (0.010)	0.025*** (0.003)
Market Positioning _{fd}	-0.091*** (0.005)	-0.072*** (0.005)	-0.068*** (0.005)	-0.098*** (0.002)
N.Obs	264,613	128,390	56,696	264,613
Adj. R ²	0.34	0.31	0.11	-
Firm FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes

Notes: Table reports regressions on the probability of exporting the WCPV. Data are for 2006. ^a In columns (2) and (3) marginal products and destinations, those counting less than 1% and 5% respectively of total export, are removed from the sample. ^b In column (4) we estimate a Probit model with Correlated Random Effects CRE discussed in Papke and Wooldridge (1996). Standard error clustered at firm and country level.

***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

determines an increase of 2 and 3 percentage point in the probability of exporting the WCPV_f.

As before the market positioning variable has the strongest effect on a firm's product vector diversification. Indeed, the closer the position of a firm's with respect to its competitors in a destination, the more likely that this firm diversifies its product vector by exporting goods different from its core. Doubling the market positioning variable determines a decrease in the probability of exporting a firm's WCPV of 9 percentage points in both Italy and France.

In columns 2-4, we show that this effect of the three firm-destination level variables on firms' export product vector choices is not driven by the over-inflation of zeros associated to "irrelevant" products or countries nor by the dichotomic nature of the dependent variable. The coefficients on the two explanatory variables are significant well beyond the 1% threshold throughout all our different specifications.

5.3 How different is a firm's local product vector from its core vector?

The binary measure $Dwcpv_{fd}$ used as dependent variable in the previous regression is not informative about “how much” different two products vectors are. Consider the following illustrative example

$$CPV_f = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \quad LPV_{f1} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad LPV_{f2} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} .$$

It is apparent that both LPV_{f1} and LPV_{f2} are different from the CPV_f but the simple dummy variable does not tell us to what extent they differ. In order to measure the distance from the CPV_f we adopt the Levenshtein (or Edit) distance, a string metric developed by sequence analysis for measuring the difference between two sequences. It represents the minimum number of single edits (insertion, deletion, substitution) required to change one sequence into the other divided by the number of elements of the longest one.³¹

In the previous example the $Lev(CPV_f, LPV_{f1})$ is equal to 0.75: 3 changes are needed to transform LPV_{f1} into the CPV_f and the length of the longest sequence is 4. Instead, the $Lev(CPV_f, LPV_{f2})$ is equal to 0.25: 1 change is required to transform LPV_{f2} into the CPV_f and the length of the longest sequence is 4. More generally, in our example any additional difference between the two sequence affects the Levenshtein distance by 1/4.

The Levenshtein distance is thus computed between each local and core product vectors for all the firms in our sample. The firms' average distance is reported in the last row of each panel of Table 4. We observe remarkable degree of heterogeneity among a firm's LPV_{fd} and the corresponding CPV_f . Indeed, the observed average distance is 0.44 both in Italy and France and the values span the whole theoretical range of the Levenshtein distance (0,1). With an average number of 12 products exported this implies that 5.4 changes are needed to transform a local into a core product vector.

To investigate factors behind the observed degree of heterogeneity among LPV_{fd} in terms of Levenshtein distance from the CPV_f , we propose the following regression model

$$LevD_{fd} = \beta X_{fd} + \alpha_f + \alpha_d + \epsilon_{fd}$$

where we regress $LevD_{fd}$, the Levenshtein distance of the LPV exported by firm f in country d from the corresponding CPV , on the three firm-country level variables contained in X_{fd} , firm and destination fixed effects.

Results of the regressions, reported in Table 6, turn out to be in line with the previous ones: the higher the demand and the level of competitive pressure the lower the degree of heterogeneity among LPVs. They all look more like the CPV. On the contrary, the closer a firm's positioning

³¹See Appendix C for more details on this measure.

Table 6: REGRESSIONS ON THE NORMALIZED LEVENSHTTEIN DISTANCE FROM CPV

Dependent variable:					
	Benchmark (1a)	(1b) ^a	1% cut (2) ^b	5% cut (3) ^b	Fractional model with CRE (4) ^c
<i>ITALY</i>					
Demand level _{fd}	-0.001*** (0.0004)	-0.011*** (0.004)	-0.006*** (0.001)	-0.007*** (0.002)	-0.003*** (0.0004)
Competition _{fd}	-0.004*** (0.001)	-0.008*** (0.003)	-0.019*** (0.003)	-0.018*** (0.005)	0.0003 (0.001)
Market Positioning _{fd}	0.015*** (0.001)	0.106*** (0.004)	0.039*** (0.002)	0.061*** (0.003)	0.019*** (0.0003)
N.Obs	621,046	621,046	279,410	98,641	621,046
Adj. R ²	0.70	0.70	0.50	0.41	-
<i>FRANCE</i>					
Demand level _{fd}	-0.001* (0.0007)	-0.015* (0.008)	-0.004** (0.002)	-0.006* (0.003)	-0.006*** (0.001)
Competition _{fd}	-0.005*** (0.001)	-0.011*** (0.003)	-0.023*** (0.004)	-0.034*** (0.006)	-0.008*** (0.001)
Market Positioning _{fd}	0.014*** (0.001)	0.094*** (0.005)	0.033*** (0.002)	0.044*** (0.003)	0.017*** (0.001)
N.Obs	245,006	245,006	110,247	41,949	245,006
Adj. R ²	0.66	0.66	0.45	0.40	-
Firm FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes

Notes: Table reports regressions on the normalized Levenshtein distance from the CPV. Data are for 2006. ^a Column (1b) reports standardized coefficients. ^b In columns (2) and (3) marginal products and destinations, those counting less than 1% and 5% respectively of total export, are removed from the sample. ^c In column (4) we estimate a Probit model with Correlated Random Effects CRE discussed in Papke and Wooldridge (1996). Standard error clustered at firm and country level.

***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

with respect to its competitors, the higher the Levenshtein distance from the CPV.

These results complement those obtained using export shares confirming that firms are active in adjusting their product portfolios to demand, competition level and market positioning in each destination they serve. Standardized coefficients in column 1b reveal that the ranking of the three firm-country level variables remains equal to that observed before, that is the most important variable offering an explanation for firms' variation in product mix across destinations is the market positioning, followed by demand heterogeneity and competition level.

As before, in columns 2 and 3 we drop marginal products and countries, defined as those involving less than 1% or 5% of the overall exports of each firm. Removing such transactions might make the identification cleaner. The coefficients on both dependent variables are preserved. Finally, we implement a fractional probit model to account for the nature of the dependent variable.

Results, reported in column 4, are essentially unchanged.

6 Conclusion

Although recently an increasing number of papers has recognized the preeminence and the role of multi-product firms in international trade, it appears that very little is known about the behaviour of these firms on foreign markets. How do firms diversify their product portfolio across destinations? What are the main factors driving the variation of a firm's product mix choice across markets?

Existing theoretical models of multi-product and multi-country firms predict a perfectly stable hierarchy in the products exported by a firm across destinations. Given that firms face declining efficiency in supplying their less successful products away from their core competency, they should enter export markets with the most efficient product first and then expand their scope moving-up the marginal-cost ladder. This would imply a high level of correlation of a firm's product sales across its destinations.

This paper claims that firms exhibit a more fickle product mix across countries. Using information for the universe of Italian and French manufacturing exporters, we establish new stylized facts consistent with a more complex model where firms do not follow a rigid ordering in their product mix across markets but rather they adapt their choices to better match with country characteristics.

First, in studying product hierarchies across countries we move beyond simple rank correlations and we use a dissimilarity index based on product market shares which allows us to better capture the economic importance of each product in terms of export shares. The analysis confirms that there is a substantial deal of dissimilarity between local product share vectors of the same firms across destinations. Moreover, by using sequences of product names we provide new insights on the extent a firm's products portfolio changes across destinations. Again, the empirical evidence suggests that firms exhibit a substantial departure from an equable product mix composition across markets. All these results resist when we control for a firm's choice of not exporting an available product to a given destination, an explicit likely to contain relevant information.

Second, the paper offers possible explanations for the observed evidence. Demand heterogeneity, aiming to capture asymmetric preferences for different varieties, market competition and market positioning emerge as three significant factors in explaining the variety-country variability in the export patterns of multi-product firms.

Overall, these empirical findings paint a more complex picture of the behaviour of multi-product firms in foreign markets. Given the significance of these firms in international trade, we believe it is important to understand their export patterns. An important avenue for future research is understanding how the adaptation mechanism across destinations emerged in our empirical analysis affects the welfare and distributional consequences of international trade.

References

- ARKOLAKIS, C. AND M.-A. MUENDLER (2010): “The Extensive Margin of Exporting Products: A Firm-level Analysis,” NBER Working Papers 16641, National Bureau of Economic Research, Inc.
- (2013): “Exporters and their products: a collection of empirical regularities,” *CESifo Economic Studies*, 59, 223–248.
- BERNARD, A. B., J. B. JENSEN, AND P. K. SCHOTT (2009): “Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Goods,” in *Producer Dynamics: New Evidence from Micro Data*, National Bureau of Economic Research, Inc, NBER Chapters, 513–552.
- BERNARD, A. B., A. MOXNES, AND K. H. ULLTVEIT-MOE (2014): “Two-Sided Heterogeneity and Trade,” NBER Working Papers 20136, National Bureau of Economic Research, Inc.
- BERNARD, A. B., S. J. REDDING, AND P. K. SCHOTT (2011): “Multi-Product Firms and Trade Liberalization,” *Quarterly Journal of Economics*, forthcoming.
- BRAY, J. R. AND J. T. CURTIS (1957): “An ordination of upland forest communities of southern Wisconsin,” *Ecological Monographs*, 27, 325–349.
- DI COMITE, F., J.-F. THISSE, AND H. VANDENBUSSCHE (2014): “Verti-zontal Differentiation in Monopolistic Competition,” *Journal of International Economics*, forthcoming.
- EATON, J., S. KORTUM, AND F. KRAMARZ (2011): “An Anatomy of International Trade: Evidence From French Firms,” *Econometrica*, 79, 1453–1498.
- ECKEL, C., L. IACOVONE, B. JAVORCIK, AND J. P. NEARY (2011): “Multi-Product Firms at Home and Away: Cost- versus Quality-based Competence,” CEPR Discussion Papers 8186, C.E.P.R. Discussion Papers.
- ECKEL, C. AND J. P. NEARY (2010): “Multi-Product Firms and Flexible Manufacturing in the Global Economy,” *Review of Economic Studies*, 77, 188–217.
- GAULIER, G. AND S. ZIGNAGO (2010): “BACI: International Trade Database at the Product-Level. The 1994-2007 Version,” Working Papers 2010-23, CEPII research center.
- KEE, H. L. AND K. KRISHNA (2008): “Firm-Level Heterogeneous Productivity and Demand Shocks: Evidence from Bangladesh,” *American Economic Review*, 98, 457–62.
- MAYER, T., M. J. MELITZ, AND G. I. P. OTTAVIANO (2014): “Market Size, Competition, and the Product Mix of Exporters,” *American Economic Review*, 104, 495–536.
- MUNCH, J. R. AND D. X. (2008): “Decomposing Firm-level Sales Variation,” EPRU Working Paper Series 2009-05, Economic Policy Research Unit (EPRU), University of Copenhagen. Department of Economics.

PAPKE, L. E. AND J. M. WOOLDRIDGE (1996): “Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates,” *Journal of Applied Econometrics*, 11, 619–32.

Appendix A

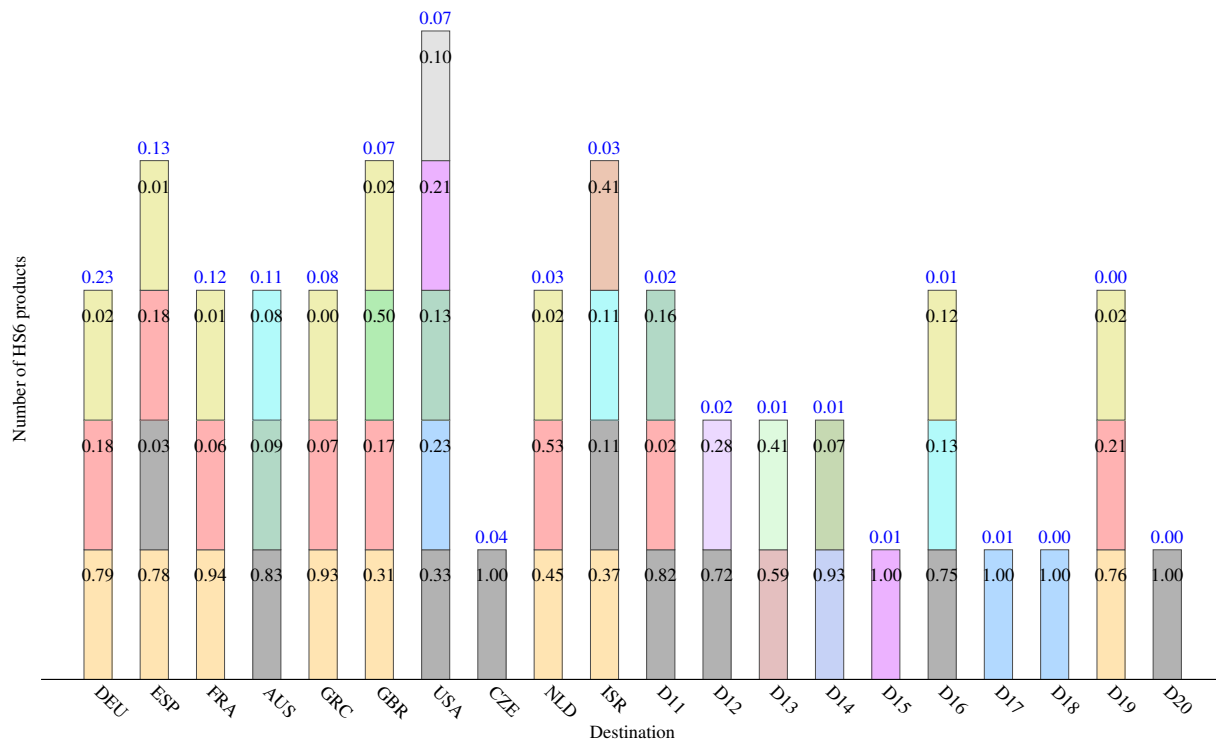


Figure A1: The figure shows an illustrative example of an Italian firms producing electrical motors and shipping 16 different products in 20 different destinations.

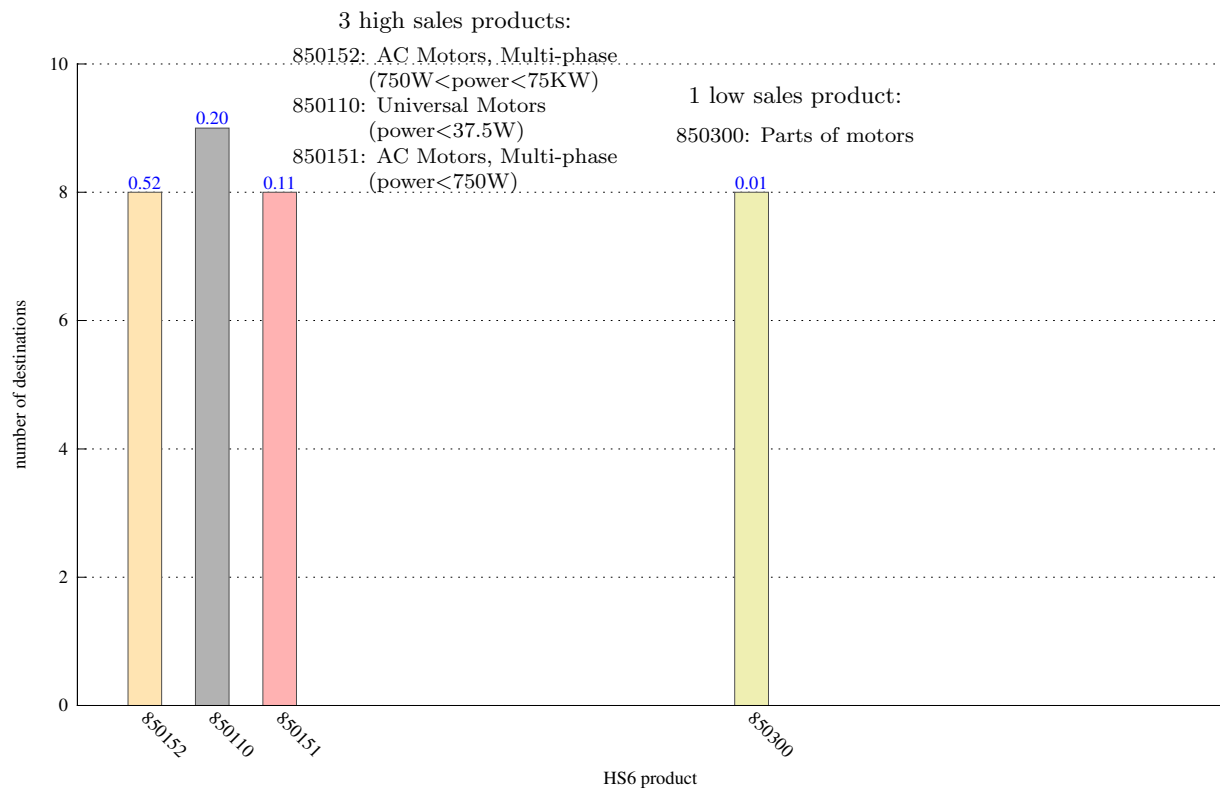


Figure A2: The figure shows an illustrative example of an Italian firms producing electrical motors and shipping 16 different products in 20 different destinations.

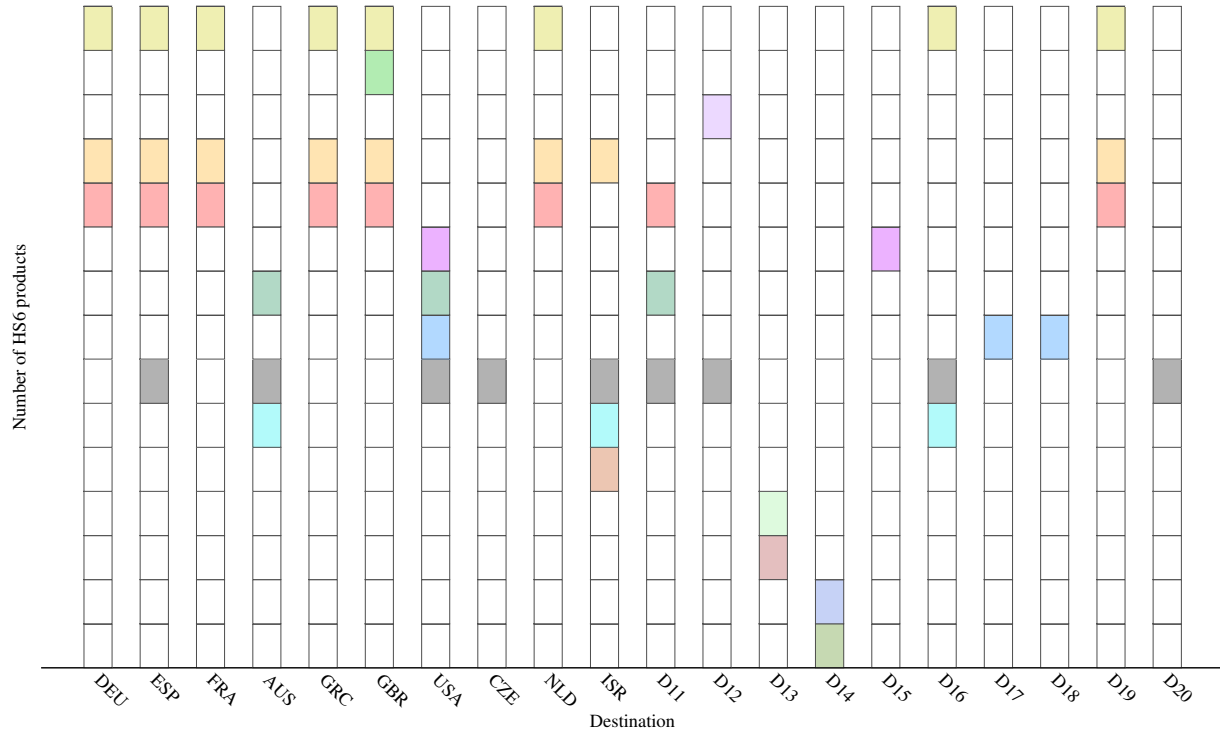


Figure A3: The figure shows the *Local product vector* (LPV_{fd}) for the electric motors producers. Within each destinations there are some products exported (colored boxes) and some other products that are not exported (white boxes).

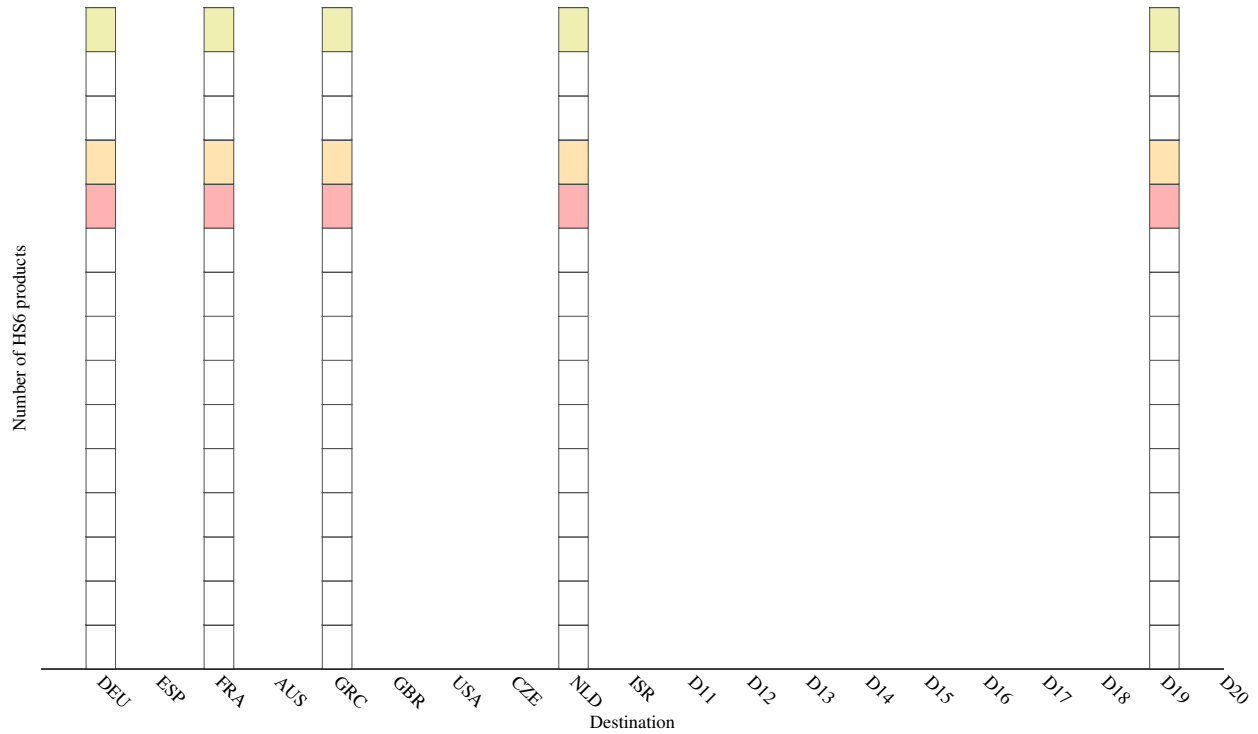


Figure A4: The figure shows the *Core product vector* (CPV_f) for the electric motors producers. This is the most frequent arrangement observed for firm f across destinations.

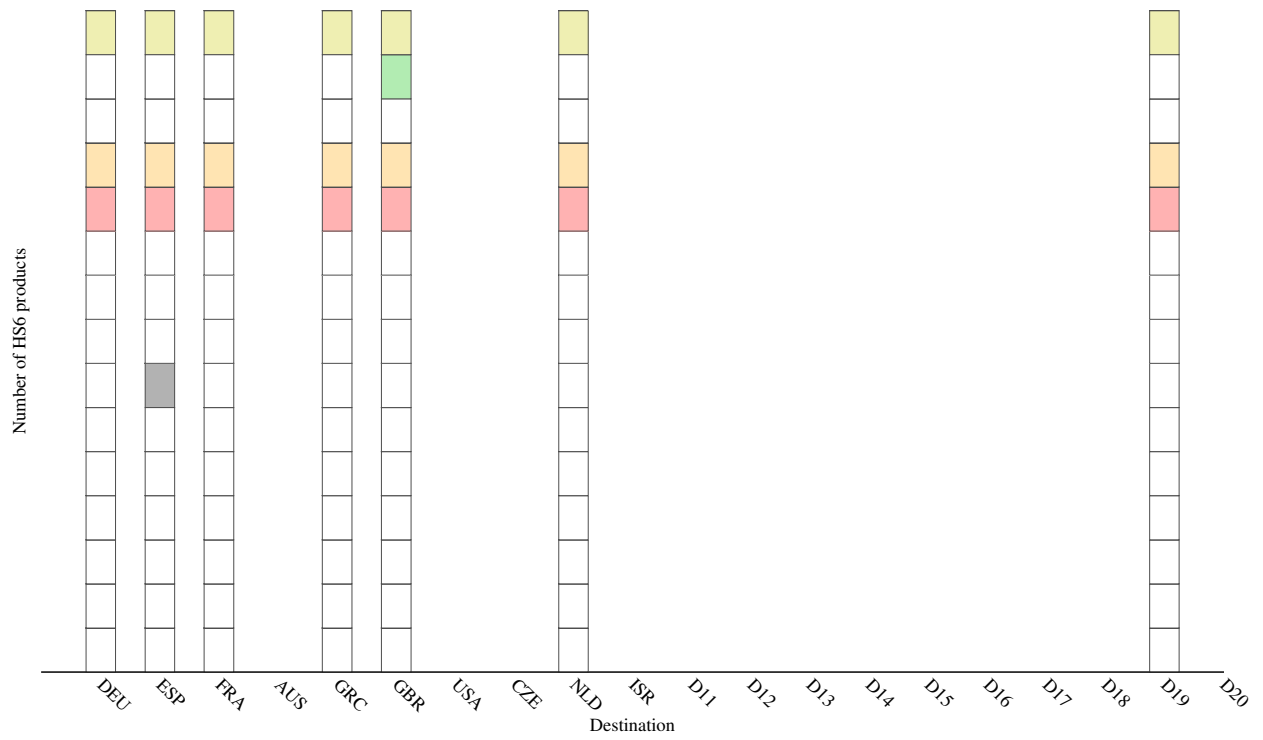


Figure A5: The figure shows the *Weak core product vector* ($WCPV_f$) for the electric motors producers. This is any arrangement which leaves unchanged all the cells with 1 in the CPV_f .

Appendix B

Exports and Number of exporting firms: share by type of firms, 2006

	Manufacturers (1)	Wholesalers & Retailers (2)	Others (3)	Total (4)
ITALY				
Total exports (billion euros)	271.1	38.5	9.1	318.7
# Firms	74,365	48,643	14,861	134,579
FRANCE				
Total exports (billion euros)	268.7	79.8	28.3	376.9
# Firms	32,432	44,379	25,473	102,284

Notes: Table reports descriptive statistics on the Italian and French dataset for 2006.

Appendix C

6.1 Bray-Curtis dissimilarity index

The Bray-Curtis dissimilarity index, also known as Sørensen index, is a well-known way of quantifying the difference between samples.³² Formally, the Bray-Curtis dissimilarity index between vector i and vector j is defined as

$$BC_{i,j} = \frac{\sum_{k=1}^K |i_k - j_k|}{\sum_{k=1}^K |i_k + j_k|} \quad , \quad (5)$$

where i_k and j_k represent the number of elements observed in vector i and j along the k th dimension. The Bray-Curtis index is symmetric and it ranges from 0, when the two vectors are identical, to 1 where the two vectors are disjoint.³³ The BC measure is not a euclidean distance since it does not satisfy the triangular inequality axiom.

Usually, the BC index is computed on count data.³⁴ Consider the two following vectors i and j with dimension $K = 3$

$$i = \begin{pmatrix} 11 \\ 0 \\ 7 \end{pmatrix} \quad j = \begin{pmatrix} 24 \\ 37 \\ 5 \end{pmatrix} \quad ;$$

the $BC_{i,j}$ between the two is 0.619 is obtained as follow

$$BC_{i,j} = \frac{|11 - 24| + |0 - 37| + |7 - 5|}{18 + 66} = 0.619 \quad . \quad (6)$$

When computed with raw count data, the BC dissimilarity index captures differences associated with both the *size* and the *shape* of the two vectors where the former regards differences in the total number of elements, $\sum_k i_k$ and $\sum_k j_k$, in vector i and j respectively while *shape* concerns the distribution of elements along different dimensions of the two vectors.

In fact, in our investigations we compute the BC dissimilarity index between a firm's GPSV and its LPSVs where the elements of the vectors are product export shares. In this case, the denominator in equation (5) is always equal to 2, and the index captures differences only in *shape*. Using the BC index is particularly useful when we observe firms exporting in different destinations products with the same ranking but with highly different export shares. Consider a simplified case

³²Originally this measure has been developed to study species abundance in different locations in ecological studies. (Bray and Curtis, 1957)

³³Two vectors i and j are disjoint if whenever there is a non-zero entry in i , there is a zero entry in j and the other way round.

³⁴You might think the number of firms in a location or number of products exported to a destination.

in which a firm's export shares vectors are

$$GPSV_f = \begin{pmatrix} 0.98 \\ 0.02 \end{pmatrix} \quad LPSV_{f,1} = \begin{pmatrix} 0.99 \\ 0.01 \end{pmatrix} \quad LPSV_{f,2} = \begin{pmatrix} 0.51 \\ 0.49 \end{pmatrix} .$$

It is straightforward to note that in this situation rank correlations between GPSV and the two LPSVs would be both 1 and that they would provide a rather imprecise assessment of the diversity of local vectors with respect to the global one.